Econ Polit (2016) 33:433–461
DOI 10.1007/s40888-016-0040-4

# The disconnection between privacy notices and information disclosure: an online experiment

**Nuria Rodríguez-Priego[1] · René van Bavel[1] · Shara Monteleone[1]**

Received: 9 March 2016 / Accepted: 17 September 2016 / Published online: 4 October 2016


**Abstract** We studied whether changes to the online environment, i.e. *nudges*, can lead to changes in privacy behaviour through an on-line experiment (n = 3229) across four European countries. The output measures were obtained through the answers to a questionnaire following a mock online exercise: one revealed the amount of personal information participants were willing to disclose, and the other whether they noticed a privacy policy link. The nudges appeared as changes in the design of a mock search engine (e.g. including an anthropomorphic character, highlighting prior browsing history or changing the look-and-feel to convey greater informality). The nudges did not lead to differences in the amount of personal information disclosed, but did affect whether participants noticed the privacy link or not. Socio-demographic factors were relevant. Compared to younger participants, older participants were less likely to reveal personal information but more likely to notice the privacy policy link. Men were more likely to reveal personal information than women, and more likely to notice the privacy policy link. Finally, significant differences were found between all countries. Participants from Italy chose to reveal least personal information (followed by those in Poland, Germany and the UK), and participants from the UK were significantly less likely to notice the privacy policy link. The implications for policy are that disclosure of personal information is resilient to small changes in the web environment, but this is not the case for awareness of a privacy policy link. Moreover, the fact that age, gender, and country of residence are relevant suggests that differentiated policy approaches depending on the target population may be warranted.

**Keywords** Privacy · Data disclosure · Nudge · Data protection · Behavioural economics

✉ René van Bavel
rene.van-bavel@ec.europa.eu

[1] Joint Research Centre, European Commission, C/Inca Garcilaso 3, 41092 Seville, Spain

**JEL Classification** D03 · D83

## 1 Introduction

Concern over privacy in the age of information is as old as the Internet itself, and has grown in tandem with the suffusion of digital technology in everyday life (Federal Trade Commission 1998). On one hand, it has become increasingly easy to disclose information about ourselves, either knowingly or unknowingly. On the other, advances in digital technology has meant that such information can be stored ad infinitum and can be processed to draw significant inferences about people (Acquisti et al. 2015).

At first glance, there is nothing wrong with disclosing personal information online. This can be understood as well-thought out, indeed 'rational', decision by people willing to surrender some privacy in exchange for (oftentimes free) access to websites and high quality, personalized services (Wu et al. 2012).

However, even though they might be aware of this trade-off, individuals do end up making disclosure decisions that they later regret (Lusoli et al. 2012; Wang et al. 2013). They might not be aware of the amount of information they are revealing, or certain cognitive or structural barriers may be stopping them from managing their privacy adequately (Acquisti 2004; Solove 2012).

Privacy notices exist to prevent this situation. They provide users with information on how and for which purpose their data will be collected, used and managed. They should mitigate regret about disclosure decisions. However, the reality is that users seldom read privacy notices, even when signing agreements online (Steinfeld 2016). Moreover, people tend to over-rely on websites that display a privacy notice, as they perceive greater protection (Martin 2015). Paradoxically, users may end disclosing more personal information in sites where there is a privacy policy link (Hoofnagle and King 2008; Groom and Calo 2011).

Behavioural insights can contribute to addressing this policy issue. They have already been applied widely over the past 15 years, in a wide range of policy areas (Bogliacino et al. 2015; van Bavel et al. 2013; Lunn 2014; World Bank 2015; Executive Order No. 13707 2015;[1] Sousa Lourenço et al. 2016). Privacy is no exception (Acquisti et al. 2012, 2015; Groom and Calo 2011; Wang et al. 2013).

Online disclosure of privacy information generally happens when navigating online, a daily activity for many, and is characterized by being habitual and guided by fast thinking (Kahneman 2011). Attempts to guide or somehow influence this behaviour should therefore target automatic (i.e. System 1) thinking, as opposed to deliberate (i.e. System 2) thinking. These stand a better chance of being effective than privacy notices, which need to be read and thought about.

The evidence shows that disclosure of personal information is context-dependent and malleable. People can be very concerned about their privacy in one situation,

[1] This Executive Order, signed by President Obama, directs US federal agencies to increase the effectiveness of their programmes by leveraging behavioural science insights.

but less so in another. Moreover, privacy behaviour can be influenced by certain features of the online environment, often by suppressing privacy concerns (Acquisti et al. 2015; Bansal et al. 2016). There is scope, therefore, for interventions that generate greater awareness of the risks involved and allow for a more circumspect disclosure of personal information.

In this paper, we explore whether people can be nudged in this direction through subtle peripheral cues in a website's design (Thaler and Sunstein 2008; Bertrand et al. 2010). In an online environment, a website's design, as well as warnings and defaults, is considered part of the *choice architecture* (Sunstein 2014). These nudges are not meant to replace privacy notices, but complement them with unobtrusive mechanisms to protect privacy.

We do not propose that nudging people away from personal data disclosure is necessarily a good thing and should be adopted as a policy objective. However, we do wish to generate greater knowledge on nudges and their effect on disclosure behaviour. This can help identify problematic practices by companies in an online environment, which can be eventually be monitored or controlled if need be. And, given that efforts to control people's disclosure of personal information is likely to remain a sensitive issue, a libertarian paternalistic approach which encourages a cautious willingness to disclose personal information while preserving people's freedom will probably be more palatable to industry and the wider public than any form of regulation.

The paper continues as follows. Section 2 presents the rationale for the study's design (including the relevant behavioural insights) and the working hypotheses. Section 3 describes the experimental protocol and treatments. Section 4 presents the results, and Sect. 5 offers a concluding discussion of the results, including limitations and policy implications.

## 2 Background and hypotheses

In our study, we tested how web design affected users' direct disclosure of personal behaviour. Such an approach is not entirely novel, as the extant research on privacy nudges (Acquisti 2009, 2010; Acquisti et al. 2015; John et al. 2009; Wang et al. 2013) and visceral notices (Calo 2012; Groom and Calo 2011) shows.[2] Data was collected through an online experiment in Germany, Italy, Poland and the UK (n = 3229).

Participants were recruited to test a new online search engine, and were randomly assigned to one of eight experimental conditions. Each condition had a subtle difference in their web design. The interaction with the website yielded data on participants' behaviour, including a measure of their passive (i.e., unwitting) disclosure of personal information. This article, however, focuses on participants' voluntary direct disclosure of sensitive information about themselves when asked a series of questions following the mock search engine test.

The experimental design took inspiration from Groom and Calo (2011). They conducted a between-participant experimental study (n = 120) using a mock search

[2] See Calo (2014) for a detailed discussion of the difference between a code, a nudge and a notice in privacy behaviour.

engine, and tested the effect of a privacy policy link, a simplified privacy notice, and a five visceral notices. These included anthropomorphic characters, a change in the look-and-feel of the website to make it more informal, and the presence of either the user's click history or their IP address. All of these conditions were replicated in our experiment. They showed that people seldom read privacy policies, and that adding a link to the privacy policy page does not significantly change users' attitudes or experience. Visceral notices, however, prove to be more effective in modulating consumer privacy concerns than traditional notices in some instances.

Another study with similar features was conducted by Bertrand et al. (2010), who used subtle peripheral cues in direct mail to influence behaviour in field experiment in South Africa. The letters contained offers for a loan, and differed in content, loan price and loan offer deadlines simultaneously. The changes in content were based on the literature of how frames and cues affect choices, some of which also guided our study (see information overload and anthropomorphic characters below). Results showed that showing fewer example loans (to avoid cognitive load), not suggesting a particular use for the loan, or including a photo of an attractive woman increases loan demand by about as much as a 25 % reduction in the interest rate.

In a study such as this one, it is difficult to predict ex ante which behavioural insights, translated into an experimental treatments, will affect behaviour. A central premise in psychology is that context matters, and prior findings will not necessarily carry over to the present context, namely an online experiment in a European setting (Bertrand et al. 2010). However, Groom and Calo (2011) and Bertrand et al. (2010) offer a starting point for systematic experimentation in this field. The design of our study, therefore, was guided by the following behavioural insights, which were in turn translated into working hypotheses.

### 2.1 The presence of privacy notices

As stated earlier, people might not read privacy notices (Steinfeld 2016), but their mere presence can lead people to assume that certain privacy standards are being adhered (Hoofnagle and King 2008). Privacy notices make a website appear more trustworthy, which in turn elicits greater disclosure of personal information (Groom and Calo 2011).

*Hypothesis 1* A website in which a link to privacy notice is displayed will lead to greater disclosure of personal information than a website without a privacy notice link.

### 2.2 Information overload

Another insight that applies to navigating online is the effect of information overload (Jacoby et al. 1974; Chen et al. 2009). Internet users are faced of overwhelming amounts of information that they need to make sense of. Long privacy notices compete with other sources of information for user's attention. Therefore, shorter, more succinct, privacy notices should be more effective (Groom and Calo 2011).

*Hypothesis 2* A website with a simplified privacy notice will lead to less disclosure of personal information than a website with a 'traditional' privacy notice.

### 2.3 Informality

Users' perception of a website can depend on its look-and-feel. For example, an amusing design increases trust (Robins and Holmes 2008). Also, a more frivolous or informal look, with brighter colors and less formal fonts, leads to greater disclosure of personal information (John et al. 2009).

*Hypothesis 3* A website with an informal look-and-feel will lead to greater disclosure of personal information than a website with a more formal design.

### 2.4 Anthropomorphism

The presence of anthropomorphic characters on a website is intended to evoke the feeling of being in the presence of another human being. Therefore, some features of this presence are expected to carry over to an online setting, such as increased trustworthiness and credibility (Heckman and Wobbrock 2000; Qiu and Benbasat 2009). Online social presence can also increase the feeling of being observed, which can reduce personal information disclosure (Groom and Calo 2011; Moon 2000). If an anthropomorphic character, in addition, is dynamic rather than static (i.e., it moves or speaks or has its eyes follow the cursor) the feeling of being observed should be reinforced (Bailenson et al. 2006).

*Hypothesis 4* A website that displays an anthropomorphic character will lead to less disclosure of personal information than a website where an anthropomorphic character is not displayed.

*Hypothesis 5* A website that displays a dynamic anthropomorphic character will lead to less disclosure of personal information than a website where the anthropomorphic character is static.

### 2.5 Self-focused attention

People can focus their attention on a number of things: their emotions, the task at hand, their appearance, etc. 'Self-focused attention' refers to attention directed at the aspects of the self (Bögels and Mansell 2004). It is 'public' when it refers to aspects of the self that can be judged by others (Nass et al. 1998), and it is presumed to inhibit the disclosure of personal information (Joinson and Paine 2007). In an online setting, public self-focused attention can be heightened by features such as displaying users' IP addresses or browsing histories (Groom and Calo 2011).

*Hypothesis 6* A website that displays data which can identify the user's terminal (IP address, location, and browser) will lead to less disclosure of personal information than a website where this data is not shown explicitly.

*Hypothesis 7* A website that displays the URL of each external website visited by the user during the study will lead to less disclosure of personal information than a website where this information is not displayed.

The study registered users' disclosure of personal information (see section on methodology for a description of the measure). In addition, it included participants' self-reported awareness of a link to the website's privacy policy. This measure was included to test whether the subtle peripheral cues operated via intuitive or deliberative processes (Bertrand et al. 2010; Kahneman 2011). Greater awareness of the privacy policy link, combined with a change in the amount of personal information disclosed, would suggest the presence of deliberate thinking. For all nudges described above, we hypothesized that their effect would be via automatic behaviour.

*Hypothesis 8* Effects on behaviour will not be accompanied by similar effects on the measure 'awareness of a privacy policy link'.

## 3 Experimental design

The sample consisted of 3229 participants recruited in Germany, Italy, Poland and the UK. In addition to those participants, 2727 participants started but did not complete the entire study. The experiment was translated to the four languages of the countries selected. All participants were randomly assigned to one of the seven experimental conditions or to the control group. The study targeted around 400 subjects per experimental condition in the total sample, and around 100 subjects per experimental condition in each country. The median completion time was 17.48 min and the average completion time was 23.47 min (SD: 47.80).

### 3.1 Experimental protocol

Online participants were recruited and 'passed' to a controlled server. The survey that ran on that server was coded in PHP, and survey responses were saved in an SQL database, before being translated into Excel and STATA format. Upon reaching the server, each participant was randomly assigned to one of the eight experimental conditions. Participants were recruited from four European countries (UK, Germany, Italy, Poland), and were assigned to different language versions of the survey depending on their country of provenance.

Before participating in the experiment, participants had to sign an informed consent. Participants' recruitment was subcontracted but controlled by the researchers. The subcontractor recruited panel members from different sources, including graphical and text banner placement on partners' websites (including social media, news, search, and community portals), targeted emails, co-registration offers on partners' websites, and telephone recruitment of targeted populations. Each recruitment source is routinely vetted by the subcontractor, including monitoring response quality and screening and updating demographic variables to allow for sample representativeness.

Based on our requests, the subcontractor first prepared a sample plan that focused on the goal of sampling a representative group of participants from four European countries; then, based on the sample plan, we set quotas to balance demographic variables and performed real-time quota management during the run of the study.

As noted above, all participants were redirected from the subcontractor sample to a server controlled by the researchers. On the server, they were automatically segmented by country of provenance, and randomly assigned to one of the eight versions of the survey. In order to participate in the survey, participants had to: be at least 18 years old or older; be connecting from the appropriate country, among the four countries chosen for the study; have at least 30 uninterrupted minutes to complete the study; have a reliable Internet connection; be using a desktop or laptop.

During the experiment, subjects were asked to evaluate a new search engine by searching for several pre-established questions. However, this was a pretext to observe their behaviour. This process allowed for the collection of information on the IP address of participants' computers, the web browser used and web pages that were visited, which would be relevant later on in some of the treatments (see next subsection). Participation in the experiment could not be discontinued, otherwise it would be considered invalid. The search engine was fully functional; it was a mock up from a real search engine (participants received this information at the end of the experiment).

The search engine had an ad-hoc name 'Re-Search Engine', a logo, a search box and, below, an area displaying search results. It was adapted and modified according to the needs of the seven experimental conditions or control group. It could direct participants to existing external webpages. However, it was ensured that the subjects returned to the search engine website once they had found the answers to the search queries, so that they continued with the experiment. The questions that the participants were asked were displayed above the search box. Below the search box, another box was provided in which participants could type their responses.

Finally, at the end of the experiment the software displayed separate pages, with questionnaires on Internet use and on the user interaction with the search engine. The questionnaires were also translated into all the languages of the four selected countries (Fig. 1).

### 3.2 Treatments

The seven experimental conditions and the control group were based on the design used by Groom and Calo (2011). All conditions, except the control group, included a link to a privacy notice. This would allow for testing users' willingness to read privacy policies after a treatment. The seven experimental conditions and the control group were as follows:

#### *3.2.1 Control*

The search engine did not include any privacy notice. Otherwise it displayed the same appearance as the other conditions (except for the informality condition).

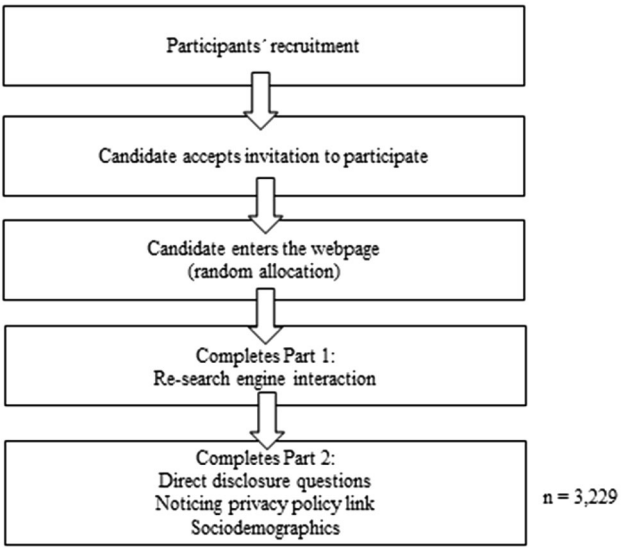



**Fig. 1** Experimental procedure flowchart

Nuances of blue or grey were used throughout the webpages to transmit authority and seriousness.

### *3.2.2 Traditional*

This experimental condition displayed a clickable privacy policy link at the top of the far-right column. Clicking the link would open a page displaying a traditional privacy notice, consisting of written text, explaining precisely what data were going to be collected by the mock search engine and how these data would be used.

### *3.2.3 Simplified*

This experimental condition displayed the same link to a privacy notice as in the traditional condition, but which led to a much simpler version of the same notice. The same information was conveyed in simpler language and with the help of a table. The rationale was to help participants to be informed of how their data would be used, managed and stored with an easy-to-read text (Groom and Calo 2011).

### *3.2.4 Static agent*

An anthropomorphic character displayed from the shoulders up which appeared in the right column in line with the search box with a text written below. Participants were able to see a static image of this agent with the words 'what would you like to search for?' beneath it. The agent had the appearance of a customer service agent. The presence of a humanoid character seems to decrease personal data disclosure (Moon 2000; Groom and Calo 2011).

#### 3.2.5 Interactive agent

Included the same anthropomorphic character and the same text as in the static agent condition, but with moving head and eyes tracking cursor movements. Compared to the previous condition, the interactive agent should reinforce the effect on data disclosure (Bailenson et al. 2006).

#### 3.2.6 Informality

The overall appearance of the search engine website was adjusted to convey a more informal and youthful look and feel than the others. It had bright yellow background with green and blue accents and red text, with softer lines in the logo and title, rounded shapes for the buttons and Comic Sans font. The content and structure were the same as the other seven conditions. This condition should have the opposite effect to the other treatments as, according to the literature, when the background of a website has a more frivolous or informal look, users tend to disclose more personal information (John et al. 2009).

#### 3.2.7 IP information

This condition displayed the participants' real IP, location and the browser that they were using on the right side of the search engine webpage. The three corresponding messages were: '*Your IP is* […]'; '*Your current location is* […]'; '*Your current browser is* […]'. This information was collected for all participants and explained in the privacy policy, though it was only visibly displayed to the participants in this condition. This treatment and the history condition aimed to increase public-self-focused attention (Bögels and Mansell 2004), making participants be thoughtful of their actions during the experiment when they realized that they were being recorded through their IP, location, browser and URL history (Joinson 2001).

#### 3.2.8 History

This condition displayed the URL of each external website visited during the search experience on the right side of the search engine webpage. This information appeared in line with the search box. When participants visited a new site, the corresponding URL appeared at the top of the list. Click-stream data were collected for all participants and this was clarified in the privacy notices, though it was visibly displayed only to the participants in this condition.

### 3.3 Output measures

The output measures were taken from prior studies looking at the same phenomenon as follows:

#### *3.3.1 Direct disclosure*

This measure was based on the replies to ten questions about socially stigmatized behaviours (see Table 5), taken from Acquisti et al. (2012). Participants had the possibility to answer positively, answer 'never' or not to answer at all at the questions. In other words, responding was optional. The behavioural measure scored between zero (if participants answered 'never' to all the ten items) and ten (if they answered positively to all the items).

#### *3.3.2 Privacy policy link awareness*

This was a binary construct, taken from Groom and Calo (2011). Subjects were asked whether they had noticed a privacy policy link in the search engine website with two possible answers ('I noticed it' or 'I didn't notice it').

Socio-demographic data were also recorded during the experiment (see next section).

## 4 Results

### 4.1 Socio-demographics

Below, we present a series of demographic statistics sorted by gender, age, education level (Table 1) and country of provenance per treatment (Table 2). The statistics confirm that, within each country, the sample of subjects exhaustively covered diverse and balanced segments of population.

### 4.2 Direct disclosure

This construct presented a Cronbach's alpha of 0.7154, with an average interitem covariance of 0.1257. We first present some descriptive statistics on direct disclosure. Table 3 shows the distribution of the answers from participants who avoided answering to any of the stigmatized questions to participants who answered to all of them.

A Poisson regression model tested the effect of the different treatments compared to the control group on direct disclosure. We decided to include also socio-demographic variables to test if any of them had an effect in the dependent variable (see Table 4).

The results reveal that there are no significant differences at 95 % level of confidence between the subjects in the control group and the rest of the treatments. It means that, contrary to what was expected, none of the treatments had any effect on the quantity of information that participants disclosed actively (H1 to H7 are not supported according to Table 4).

However, while disclosure was resilient to subtle changes in the online environment, it was susceptible to socio-demographic factors.

**Table 1** Socio-demographic distribution of the sample

| Factors | Frequency | % |
|---|---|---|
| Gender | | |
| Female | 1574 | 49 |
| Male | 1612 | 50 |
| No answer | 40 | 1 |
| Age | | |
| 18–25 | 504 | 16 |
| 26–33 | 471 | 14 |
| 34–41 | 504 | 16 |
| 42–49 | 512 | 16 |
| 50–57 | 545 | 17 |
| 58–65 | 421 | 13 |
| 66 and above | 230 | 7 |
| No answer | 39 | 1 |
| Education level | | |
| Below high school | 89 | 2 |
| Some high school | 157 | 5 |
| High school graduate | 917 | 28 |
| Technical high school | 313 | 10 |
| Some college | 469 | 15 |
| College graduate | 928 | 29 |
| Advanced degree | 261 | 8 |
| No answer | 92 | 3 |

The proposed regression model shows significant differences between the countries having Germany as the baseline. At the top, participants in the UK are the ones that answer positively more frequently to the stigmatized behaviours. In the bottom, subjects from Italy are more cautious and avoid disclosing information within the stigmatized behaviours considered. Table 5 provides further information on median, non-response and zeros per country.

There is a gender effect on the number of items answered. A greater number of female participants did not answer or answered that they had never performed any of the behaviours listed in the questionnaire ($p < 0.01$; Table 2). In item 4 (*Have you ever looked at pornographic material?*), gender differences were particularly noticeable, as 41 % of females answered 'never' or did not answer, compared to 15 % of males (see Table 6). Figure 2 provides further information on the number of stigmatized items answered by gender. While more males answered positively than females, the trend is the same. The rate of response decreased rapidly after the third question in both cases, with a slight uptick at the last question.

There is a significant effect of age on the quantity of stigmatized information revealed. A greater number of older subjects answered negatively or did not answer to any of the items ($p < 0.01$; Table 4). This would appear to confirm the commonly-held belief that young adults care less about privacy (Hoofnagle et al. 2010). Regarding education, it did not show any effect on information disclosure.

**Table 2** Country of origin per treatment

| Factors | Frequency | % |
|---|---|---|
| Country per treatment | | |
| Control | | |
| Germany | 113 | 26.97 |
| Italy | 90 | 21.48 |
| Poland | 99 | 23.63 |
| UK | 117 | 27.92 |
| Traditional | | |
| Germany | 105 | 27.27 |
| Italy | 98 | 25.45 |
| Poland | 84 | 21.82 |
| UK | 98 | 25.45 |
| Simplified | | |
| Germany | 81 | 21.09 |
| Italy | 100 | 26.04 |
| Poland | 106 | 27.60 |
| UK | 97 | 25.26 |
| Dynamic | | |
| Germany | 106 | 27.68 |
| Italy | 90 | 23.50 |
| Poland | 94 | 24.54 |
| UK | 93 | 24.28 |
| Static | | |
| Germany | 93 | 22.41 |
| Italy | 98 | 23.61 |
| Poland | 115 | 27.71 |
| UK | 109 | 26.27 |
| Informal | | |
| Germany | 109 | 26.59 |
| Italy | 99 | 24.15 |
| Poland | 109 | 26.59 |
| UK | 93 | 22.68 |
| IP | | |
| Germany | 116 | 26.67 |
| Italy | 114 | 26.21 |
| Poland | 96 | 22.07 |
| UK | 109 | 25.06 |
| History | | |
| Germany | 90 | 22.61 |
| Italy | 110 | 27.64 |
| Poland | 100 | 25.13 |
| UK | 98 | 24.62 |

Pearson $chi^2$ (21) = 20.8197, Pr = 0.470

Likelihood-ratio $chi^2$ (21) = 21.0302, Pr = 0.457

Cramér's V = 0.0464

Gamma = −0.0073, ASE = 0.017

Kendall's tau-b = −0.0059, ASE = 0.014

**Table 3** Distribution of the answers to the stigmatized questions

| Number of items answered | Frequency total | % | Frequency Gender | |
|---|---|---|---|---|
| 0 | 149 | 5.29 | Female | 103 |
| | | | Male | 39 |
| | | | No answer | 7 |
| 1 | 294 | 10.43 | Female | 181 |
| | | | Male | 108 |
| | | | No answer | 5 |
| 2 | 487 | 17.28 | Female | 248 |
| | | | Male | 236 |
| | | | No answer | 3 |
| 3 | 602 | 21.36 | Female | 282 |
| | | | Male | 315 |
| | | | No answer | 5 |
| 4 | 560 | 19.87 | Female | 263 |
| | | | Male | 294 |
| | | | No answer | 3 |
| 5 | 346 | 12.28 | Female | 130 |
| | | | Male | 214 |
| | | | No answer | 2 |
| 6 | 170 | 6.03 | Female | 67 |
| | | | Male | 102 |
| | | | No answer | 1 |
| 7 | 90 | 3.16 | Female | 28 |
| | | | Male | 62 |
| | | | No answer | 0 |
| 8 | 42 | 1.49 | Female | 15 |
| | | | Male | 27 |
| | | | No answer | 0 |
| 9 | 19 | 0.67 | Female | 5 |
| | | | Male | 14 |
| | | | No answer | 0 |
| 10 | 59 | 2.09 | Female | 21 |
| | | | Male | 36 |
| | | | No answer | 2 |

### 4.3 Privacy policy link awareness

A logit regression model tested the effect of the different treatments compared to the control group on privacy policy link awareness. Including country of origin, gender, education level and age as independent variables allowed for comparison with the regression model for direct disclosure (see Table 7).

**Table 4** Poisson regression for direct disclosure

| Factors | Coefficient | Std. Error | z | P > \|z\| | [95 % Conf. Interval] | |
|---|---|---|---|---|---|---|
| Treatments | | | | | | |
| Traditional | 0.0379 | 0.0411 | −0.92 | 0.357 | −0.1187 | 0.0428 |
| Simplified | 0.0401 | 0.0410 | 0.98 | 0.328 | −0.0402 | 0.1205 |
| Dynamic | 0.0432 | 0.0414 | 1.04 | 0.297 | −0.0381 | 0.1245 |
| Static | 0.0593 | 0.0406 | 1.46 | 0.144 | −0.0202 | 0.1390 |
| Informal | 0.0132 | 0.0405 | 0.32 | 0.746 | −0.0664 | 0.0927 |
| IP | 0.0154 | 0.0401 | 0.38 | 0.701 | −0.0634 | 0.0942 |
| History | 0.0458 | 0.0413 | 1.11 | 0.268 | −0.0352 | 0.1268 |
| Country | | | | | | |
| Italy | −0.1338*** | 0.0315 | −4.25 | 0.000 | −0.19552 | −0.0722 |
| Poland | 0.2063*** | 0.0295 | 7.00 | 0.000 | 0.1485423 | 0.2641 |
| UK | 0.0598** | 0.0301 | 1.99 | 0.046 | 0.0009511 | 0.1188 |
| Other | | | | | | |
| Gender | −0.2096*** | 0.0208 | −10.08 | 0.000 | −0.2504 | −0.1688 |
| Education | −0.0022 | 0.0082 | −0.27 | 0.789 | −0.0183 | 0.0139 |
| Age | −0.0236*** | 0.0057 | −4.12 | 0.000 | −0.0349 | −0.0124 |
| Constant | 1.410 | 0.0494 | 27.78 | 0.000 | 1.275 | 1.469 |

*** $p < 0.01$, ** $p < 0.05$

Baseline treatment: control

Baseline country: Germany

Gender: female = 1

Number of observations = 2726

LR $chi^2$ (13) = 255.28

Prob > $chi^2$ = 0.0000

Pseudo $R^2$ = 0.0224

Log likelihood = −5567.9929

**Table 5** Direct disclosure per country

| Country | Median [0–10] | Non-response | Answer "Never" |
|---|---|---|---|
| Germany | 3.34 | 106 | 36 |
| Italy | 2.93 | 122 | 77 |
| Poland | 4.14 | 138 | 19 |
| UK | 3.5 | 45 | 17 |

The results show that participants who visualized the dynamic anthropomorphic ($p < 0.01$), IP address ($p < 0.1$) or history ($p < 0.05$) treatments were more likely to notice the privacy policy link when compared with participants in the control group (Table 7).

**Table 6** Direct disclosure items

| Item | |
|---|---|
| 1 | Have you ever claimed to have education that you didn't actually have? |
| 2 | Have you ever pretended not to see a beggar to avoid being seen as stingy? |
| 3 | Have you ever had sex with someone who was too drunk to know what they were doing? |
| 4 | Have you ever looked at pornographic material? |
| 5 | Have you ever had sex with the current husband, wife, or partner of a friend? |
| 6 | Have you ever known about or witnessed a serious crime and failed to report it or stop it? |
| 7 | Have you ever lied about your income to someone? |
| 8 | Have you ever fantasized about having violent non consensual sex with someone? |
| 9 | Have you ever drunk so much that you got a hangover? |
| 10 | Have you ever failed to tip a waiter in a country in which tipping is customary? |

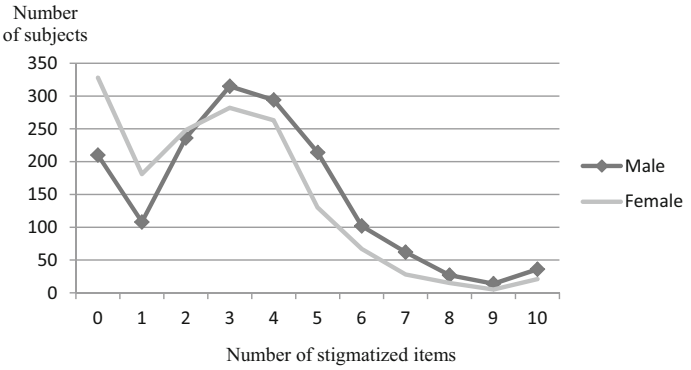



**Fig. 2** Number of stigmatized items answered by gender

The regression model shows significant differences between Germany and the UK ($p < 0.01$). When rotating the baseline country, significant differences ($p < 0.01$) emerge between the UK (92 % did not notice the link) and the other three countries (88 % in Germany, 86 % in Italy and 87 % in Poland; Fig. 3).

Males were significantly more likely to notice the privacy link compared with females ($p < 0.01$; Table 7; Fig. 4). This effect supports the claim that a privacy policy link may increase the level of disclosure (Hoofnagle and King 2008; Groom and Calo 2011).

Likewise, regarding age, younger participants were more likely to notice the privacy policy link than older ones ($p < 0.01$; Table 3), and disclosed more information on stigmatized behaviours.

However, when computing the correlation between the two behavioural measures it seems there is no connection that can relate noticing the privacy policy link and disclosing information ($r = 0.1120$, see Table 8), as can be envisaged from the regression models. Finally, it is hard to find an effect of the privacy policy link on

**Table 7** Logit regression for privacy policy link awareness

| Factors | Coefficient | Std. Error | z | P > \|z\| | [95 % Conf. Interval] | |
|---|---|---|---|---|---|---|
| Treatments | | | | | | |
| Traditional | 0.3162 | 0.2502 | 1.26 | 0.206 | −0.1741 | 0.8065 |
| Simplified | 0.3849 | 0.2467 | 1.56 | 0.119 | −0.0986 | 0.8683 |
| Dynamic | 0.5548** | 0.2396 | 2.32 | 0.021 | 0.0852 | 1.024 |
| Static | 0.3582 | 0.2451 | 1.46 | 0.144 | −0.1223 | 0.8386 |
| Informal | 0.0736 | 0.2553 | 0.29 | 0.773 | −0.4268 | 0.5740 |
| IP | 0.4589* | 0.2360 | 1.94 | 0.052 | −0.0037 | 0.9215 |
| History | 0.6056** | 0.2366 | 2.56 | 0.010 | 0.1419 | 1.069 |
| Country | | | | | | |
| Italy | 0.0444 | 0.1588 | 0.28 | 0.780 | −0.2669 | 0.3557 |
| Poland | −0.0385 | 0.1626 | −0.24 | 0.813 | −0.3572 | 0.2802 |
| UK | −0.5442*** | 0.1855 | −2.93 | 0.003 | −0.9078 | −0.1806 |
| Other | | | | | | |
| Gender | −0.4380*** | 0.1175 | −3.73 | 0.000 | −0.6683 | −0.2077 |
| Education | 0.0630 | 0.0466 | 1.35 | 0.177 | −0.0283 | 0.1544 |
| Age | −0.1172*** | 0.0325 | −3.60 | 0.000 | −0.1809 | −0.0534 |
| Constant | −1.947*** | 0.2896 | −6.73 | 0.000 | −2.515 | −1.380 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Baseline treatment: control

Baseline country: Germany

Gender: female = 1

Number of observations = 3114

LR $chi^2$ (13) = 57.33

Prob > $chi^2$ = 0.0000

Pseudo $R^2$ = 0.0264

Log likelihood = −1057.5982

direct disclosure as only three out of 3226 participants opened the notice, two of them in the *simplified* condition and one in the *IP information* condition.

## 5 Discussion and conclusions

This paper had a twofold purpose: first, it aimed to examine factors influencing users' disclosure of personal information (in particular, whether they had performed socially stigmatized behaviours in the past). For this purpose, several changes to the design of a website were tested. Country of origin, gender, education level and age were also tested as determinants of behaviour. Results show that disclosure of personal information was resilient to small changes in the online environment (i.e., the nudges did not have an effect), but that socio-demographic factors were relevant.

There was a significant difference between all the countries, with participants from the UK disclosing sensitive information the most. Men were more likely to disclose

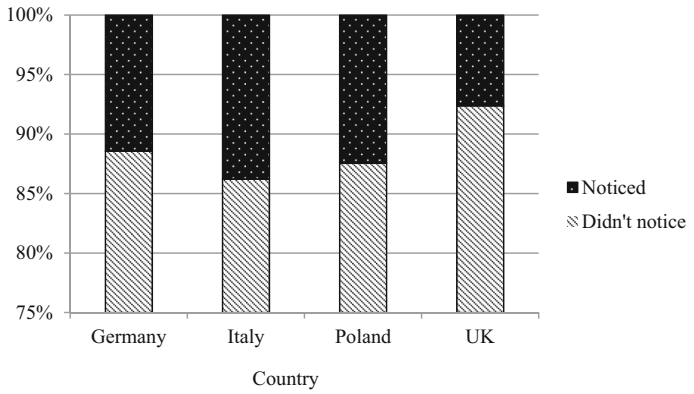



**Fig. 3** Privacy policy link awareness by country

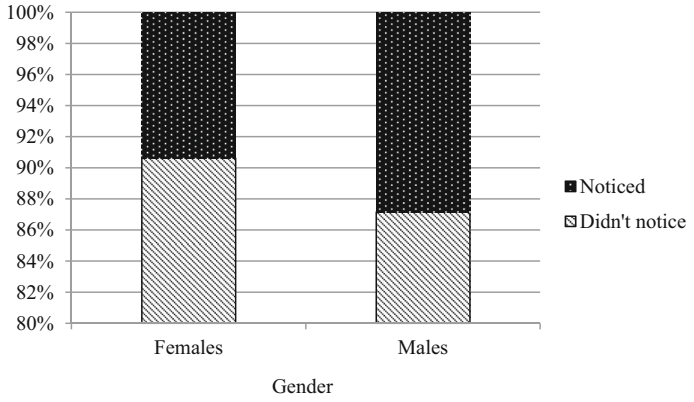



**Fig. 4** Privacy policy link awareness by gender

**Table 8** Correlations between direct disclosure and privacy policy link awareness per treatment

| Treatment | Correlation | *p* value |
|---|---|---|
| All | 0.1120 | 0.000 |
| Control | 0.1292 | 0.0132 |
| Traditional | 0.0521 | 0.3360 |
| Simplified | 0.1522 | 0.0047 |
| Dynamic | 0.1410 | 0.0101 |
| Static | 0.1861 | 0.0004 |
| Informal | −0.0706 | 0.1776 |
| IP | 0.0772 | 0.1340 |
| History | 0.2169 | 0.0001 |

than women, and younger participants more likely to disclose than older participants. This finding suggests interesting relationships between gender, age, and culture, on one hand, and information disclosure, on the other, which merit further investigation.

The second objective was to examine which factors (nudges or socio-demographics) had an effect on noticing the privacy policy link displayed in the website. Unlike with the direct disclosure measure: three nudges had an effect. A dynamic anthropomorphic character, the presence of the user's IP address, and the presence of the user's previous browsing history made it more likely for participant to notice the privacy policy link.

Regarding socio-demographics, country differences are only significant between the UK and the other three countries, but not between Germany, Italy and Poland that show a similar level of awareness. The cultural aspect seems relatively less important for this measure. However, as with direct disclosure, gender and age were relevant: women were more likely to notice the link than men, as were younger participants compared to older ones. These findings are in line with the claim that privacy notices lead to greater disclosure of personal information, perhaps due to greater perceived protection (Hoofnagle and King 2008; Groom and Calo 2011).

### 5.1 Limitations

One of the main limitations of this study involves the measure of direct disclosure. While it was based on a measure used previously in the literature (Acquisti et al. 2012), it is not without controversy. The main problem is that participants can lie, most likely by denying they have ever been involved in any socially stigmatized behaviour. They also had the option of not responding at all to the sensitive items, so behaviour was reported and not directly measured in this case. Also, findings about gender differences, with women being more cautious in their disclosure of information, might be confounded by gender bias in the questions (e.g., questions about alcohol consumption or viewing pornographic material).

If participants' replies were taken at face value, all limitations inherent to self-response exercises would apply. For this reason, we took the simple fact that participants chose to respond to these questions (which was optional) as an indicator of their willingness to engage in the exchange of sensitive information with an Internet site. This approach also has its limitations. In particular, it distinguishes between those who answer 'never' and those who do not answer. However, the fact that differences were found according to country, gender and age (albeit not experimental condition) suggests that the measure is not without its merits.

Another limitation involves the placement of the nudge in the experimental flow. It appeared while participants were searching for answers to certain questions, presumably to evaluate the effectiveness of the mock search engine. However, the questions about socially stigmatized behaviour appeared later in the process, perhaps allowing the impact of the nudge to wane.

From an policy-making perspective, the study is limited in that it only covers four European countries with an online experiment. Although every effort was made to make the environmental setting as realistic as possible, it was still an experiment, and as such participants might have had some expectations that their data would be

treated confidentially. Moreover, with an online experiment there is less control over a participant's environment (e.g., is the television on in the background; are other people in the room?). Finally, the question remains: how applicable are these findings to other EU countries not included in the sample?

### 5.2 Policy implications

In conducting an experiment on privacy nudges, this study has sought to highlight the value to policy-making of a behavioural approach to privacy. A few policy implications emerge as a result. First, small changes in the web environment do not appear to have an impact on personal data disclosure. Since absence of evidence does not imply evidence of absence, we cannot conclude that attempting to elicit changes to online behaviour through nudges is a futile exercise. It does appear, however, that nudges need to be bolder than they were in this experiment. When thinking about applying nudges as a policy tool to change behaviour, therefore, not only is it important to identify which behavioural insight might be relevant. It is fundamental to consider how the nudge will work in practice. Too subtle a nudge (as was the case here) will not have an effect. Too strong a nudge, on the other hand, might generate frustration, antagonism, and impede seamless online navigation. Finding the right balance is key.

A second implication is that nudges do affect whether participants notice a privacy link or not. This raises hope for the role of nudging in privacy. There was a tenuous link between noticing the link and disclosing more information, in line with what the literature suggests (Hoofnagle and King 2008; Groom and Calo 2011). In particular, men and older participants, both of whom were more likely to notice the privacy link, also showed a greater likelihood to disclose personal information. However, this tenuous link does not stand up to further scrutiny. In sum, noticing is a privacy policy link is malleable, but this has no significant consequences on personal data disclosure.

Thirdly, the fact that both measures are affected to a large degree (and sometimes in opposite directions) by socio-demographic factors suggests a number of cultural elements at play when it comes to disclosing personal information online. These deserve further attention suggest that policy-making in this field may need to consider differentiated approaches depending on the target population.

A final implication regards future experimentation in online privacy behaviour. While experiments such as this one are valuable for policy-making, they have their limitations (as noted above). However, the major web service providers of this world have access to vast amounts of data on their users' behaviour, much larger than anything a specific online experiment commissioned by government could ever obtain. A final recommendation, therefore, is that government work alongside with these providers and use these data to inform policy-making on privacy and data protection. Such partnerships could arrive at a series of guidelines for web interface design that allow the public to disclose personal information cautiously and conscientiously.

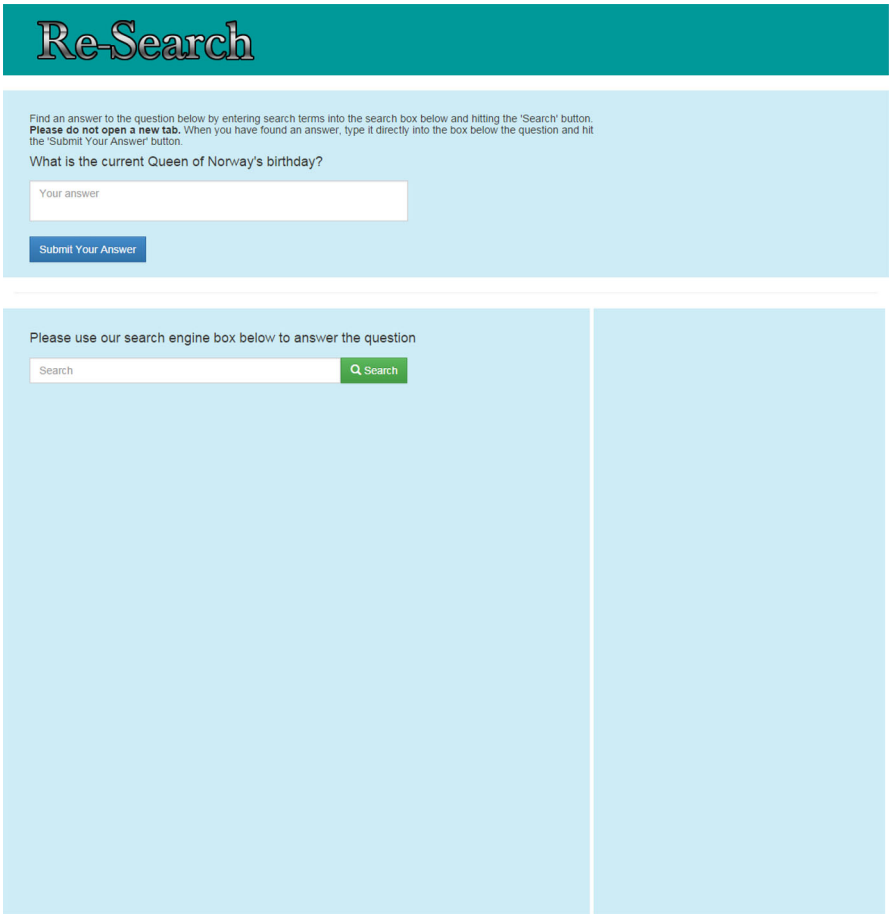
**Acknowledgments** We are grateful to Alessandro Acquisti, Norberto Andrade, Ryan Calo, Néstor Duch-Brown, Gabriele Esposito, Ioannis Maghiros and Aaron Martin for their advice and support. We also

thank two anonymous reviewers for their comments and suggestions. The views expressed in this article are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

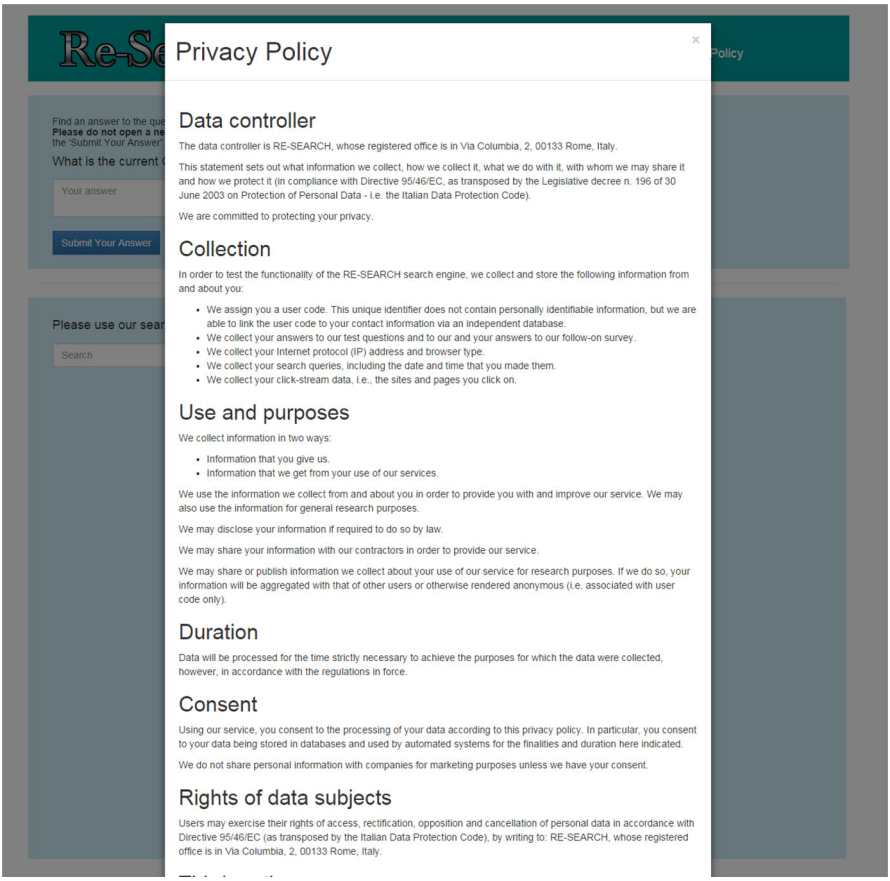


## Appendix 1: Screenshots of the different conditions

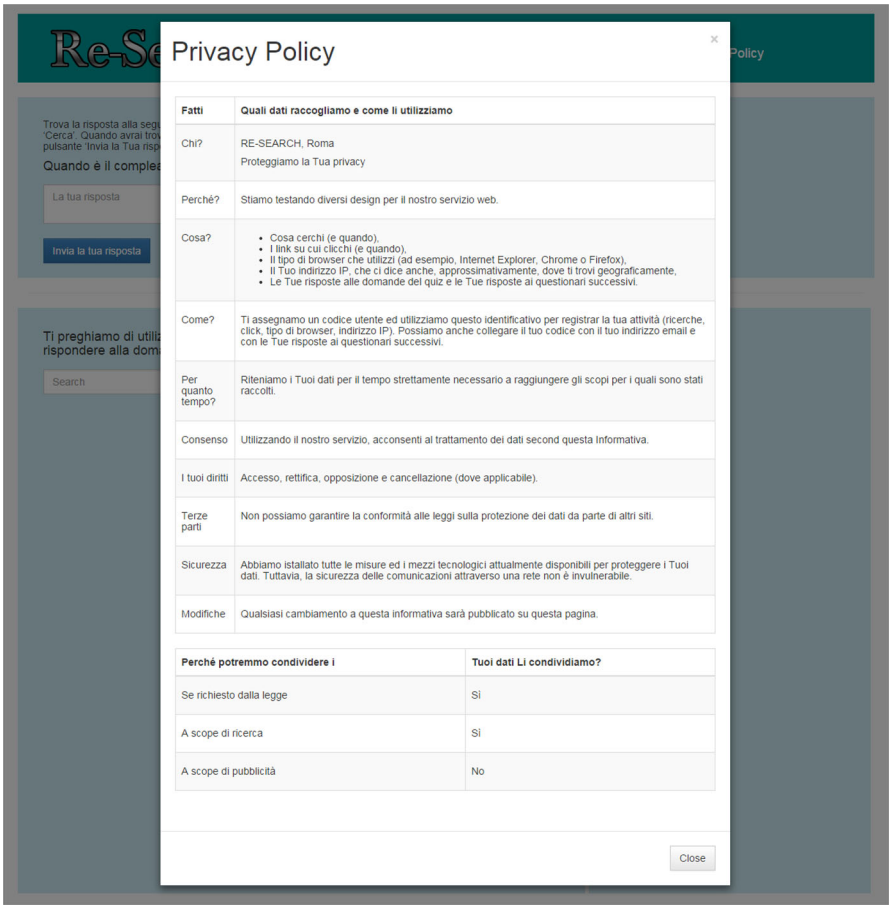
1. Control condition



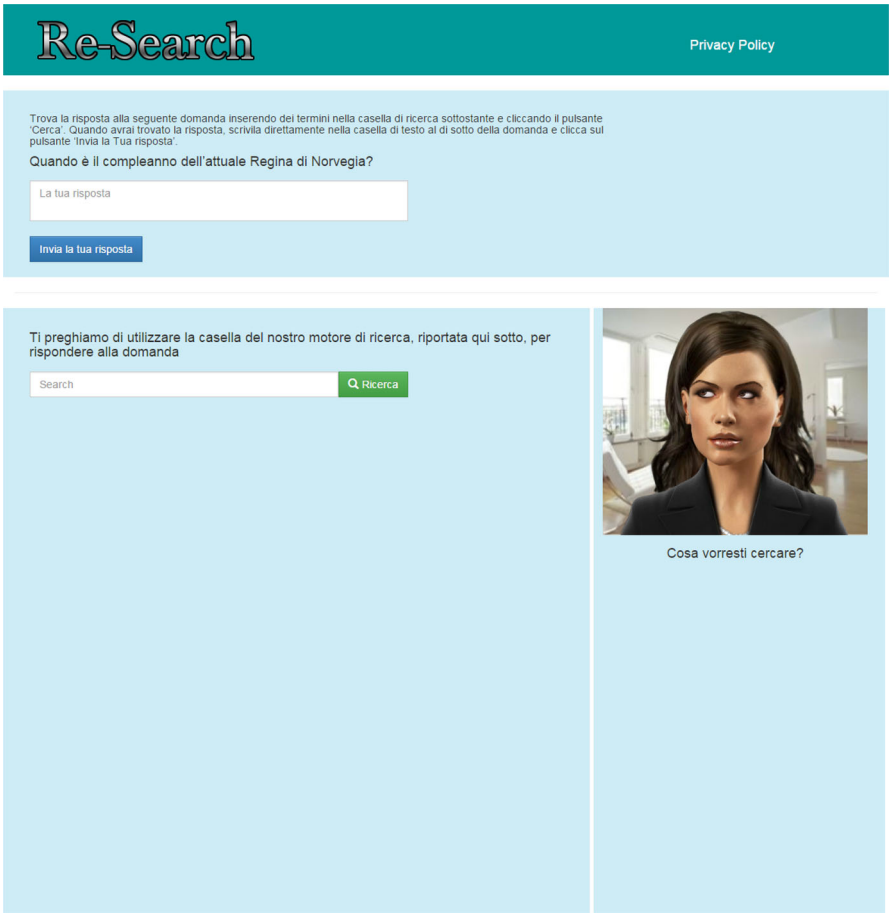
2. Traditional condition



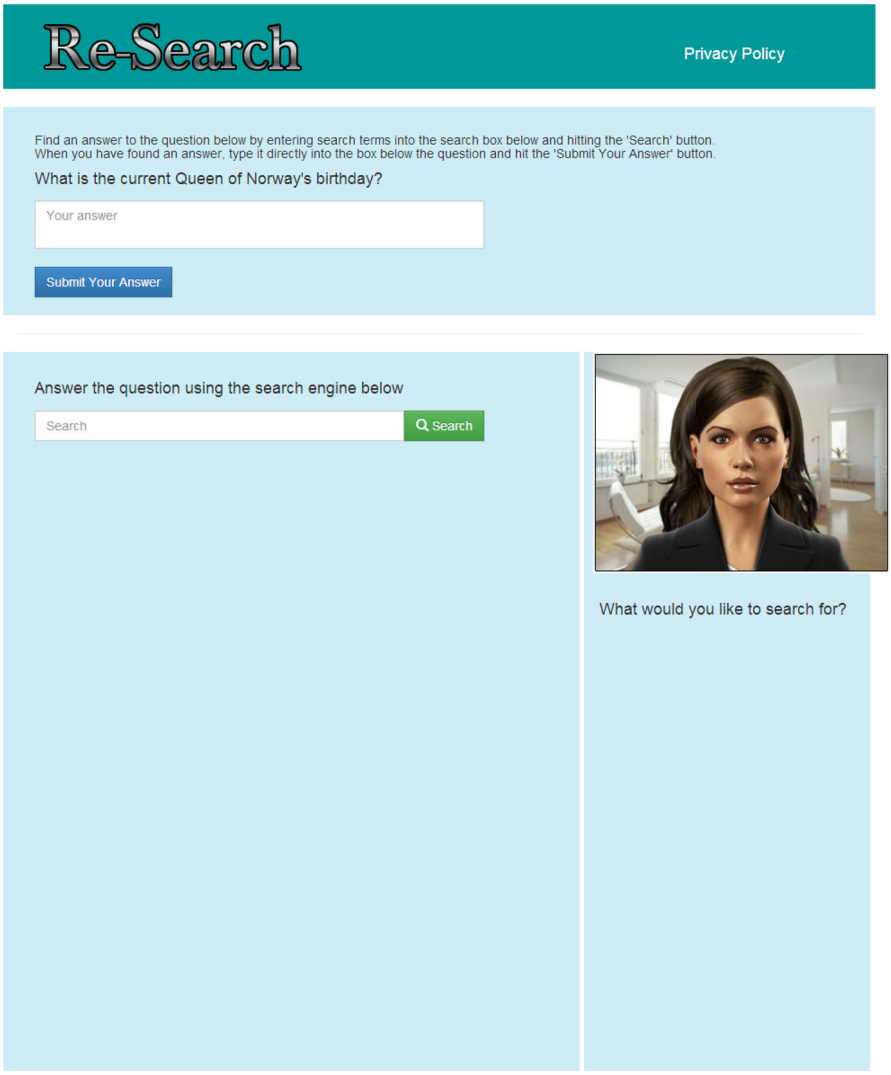
3. Simplified



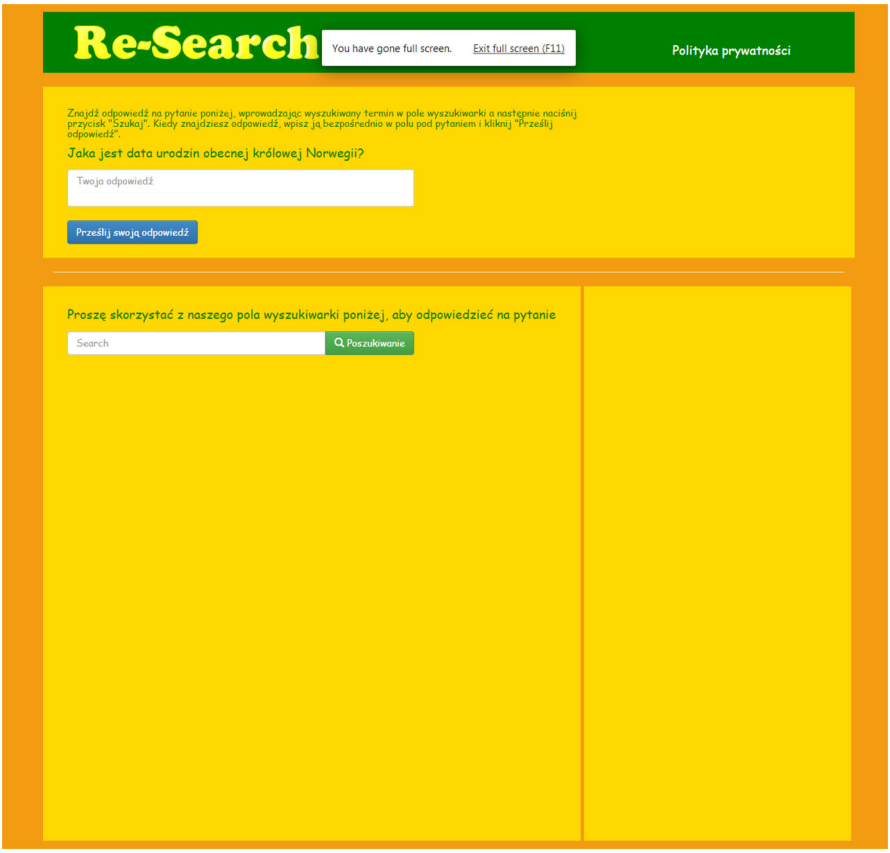
4. Anthropomorphic Dynamic



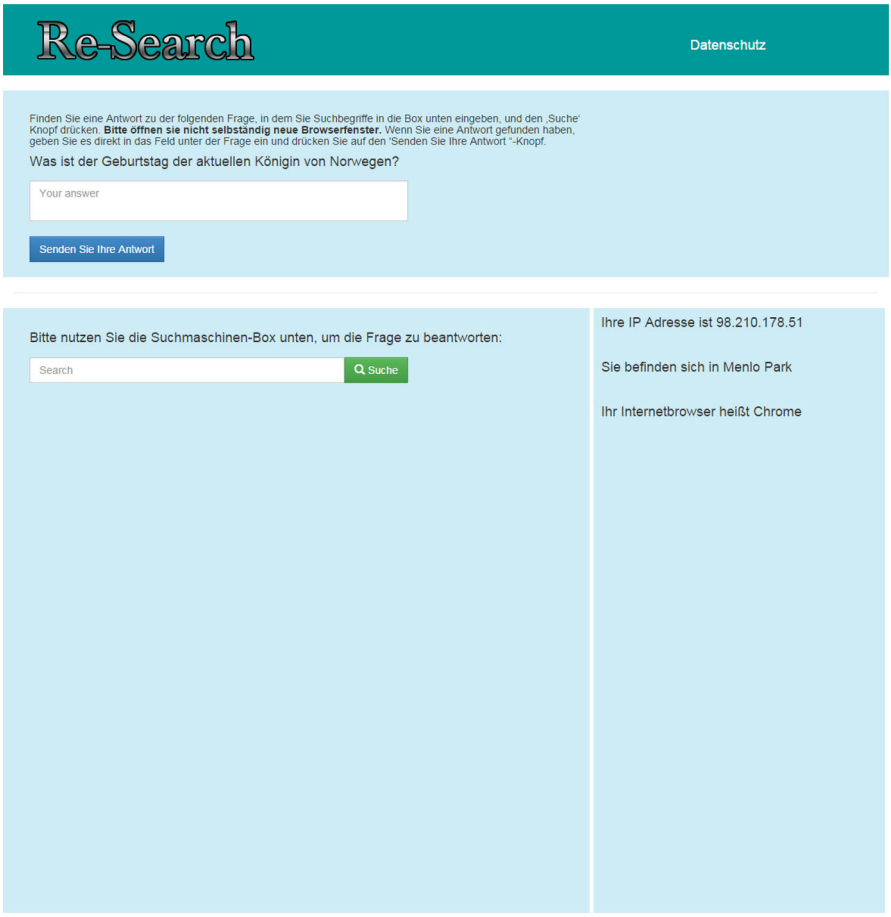
5. Anthropomorphic Static



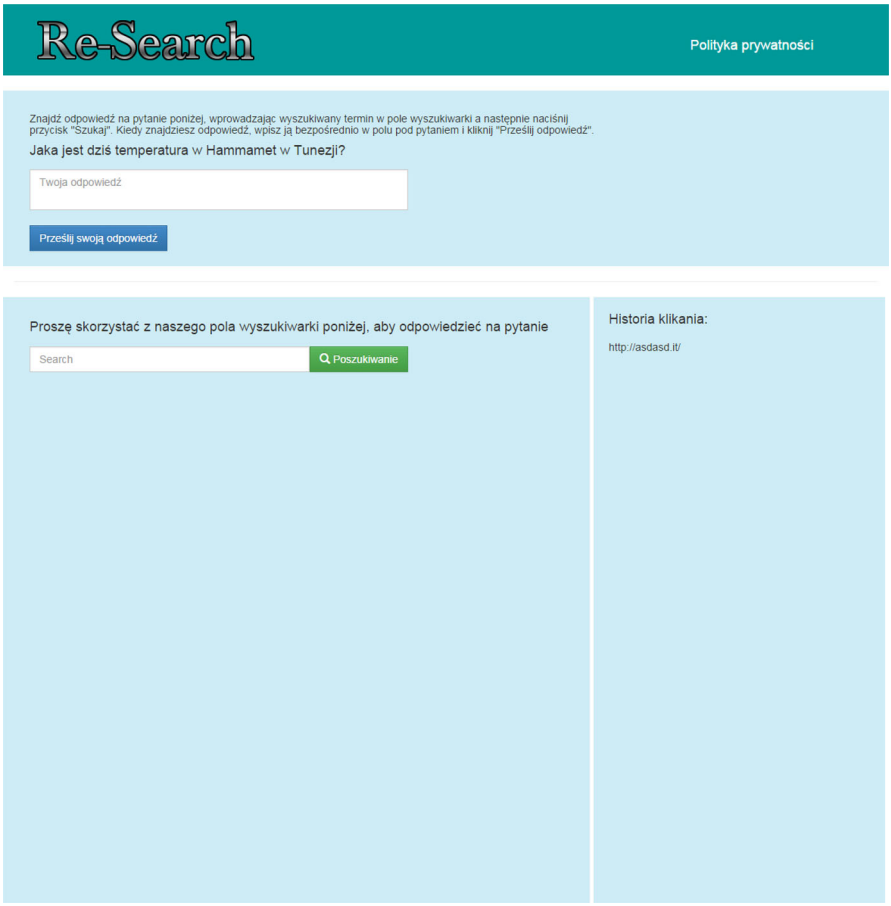
6. Informal


Re-Search
You have gone full screen.
Exit full screen (F11)
Polityka prywatności
Znajdź odpowiedź na pytanie poniżej, wprowadzając wyszukiwany termin w pole wyszukiwarki a następnie naciśnij przycisk "Szukaj". Kiedy znajdziesz odpowiedź, wpisz ją bezpośrednio w polu pod pytaniem i kliknij "Prześlij odpowiedź".
Jaka jest data urodzin obecnej królowej Norwegii?
Twoja odpowiedź
Prześlij swoją odpowiedź
Proszę skorzystać z naszego pola wyszukiwarki poniżej, aby odpowiedzieć na pytanie
Search
Poszukiwanie

7. IP


Re-Search
Datenschutz
Finden Sie eine Antwort zu der folgenden Frage, in dem Sie Suchbegriffe in die Box unten eingeben, und den ‚Suche' Knopf drücken. Bitte öffnen sie nicht selbständig neue Browserfenster. Wenn Sie eine Antwort gefunden haben, geben Sie es direkt in das Feld unter der Frage ein und drücken Sie auf den 'Senden Sie Ihre Antwort "-Knopf.
Was ist der Geburtstag der aktuellen Königin von Norwegen?
Your answer
Senden Sie Ihre Antwort
Bitte nutzen Sie die Suchmaschinen-Box unten, um die Frage zu beantworten:
Search
Suche
Ihre IP Adresse ist 98.210.178.51
Sie befinden sich in Menlo Park
Ihr Internetbrowser heißt Chrome

8. History




## References

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM conference on Electronic commerce* (pp. 21–29). New York: ACM.

Acquisti, A. (2009). Nudging privacy: behavioral economics of personal information. *Security and Privacy, 7*(6), 82–85.

Acquisti, A. (2010). From the economics to the behavioral economics of privacy: A note. In A. Kumar, D. Zhang (Eds.), *Ethics and Policy of Biometrics* (Vol. 6005, pp. 23–26). Berlin: Springer.

Acquisti, A., John, L. K., & Loewenstein, G. (2012). The impact of relative standards on the propensity to disclose. *Journal of Marketing Research, 49*(2), 160–174.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, *347*(6221), 509–514.

Bailenson, J. N., Yee, N., Merget, D., & Schroeder, R. (2006). The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators and Virtual Environments, 15*(4), 359–372.

Bansal, G., Zahedi, F. M., & Gefen, D. (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management*, *53*(1), 1–21.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., & Zinman, J. (2010). What's advertising content worth? Evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics, 125*(1), 263–306. doi:10.1162/qjec.2010.125.1.263.

Bögels, S. M., & Mansell, W. (2004). Attention processes in the maintenance and treatment of social phobia: hypervigilance, avoidance and self-focused attention. *Clinical Psychology Review, 24*(7), 827–856.

Bogliacino, F., Codagnone, C., & Veltri, G. A. (2015). The behavioural turn in consumer policy: perspectives and clarifications. *Intereconomics: Review of European Economic Policy, 50*(2), 108–114.

Calo, R. (2012). Against Notice Skepticism in privacy and elsewhere. *Notre Dame Law Review*, *87*(3), 1027–1072.

Calo, R. (2014). Code, Nudge or Notice? *Iowa Law Review, vol. 99, no. 2; University of Washington School of Law Research Paper No. 2013-04,* pp. 773–802. http://ssrn.com/abstract=2217013. Accessed 27 Sept 2016.

Chen, Y. C., Shang, R. A., & Kao, C. Y. (2009). The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment. *Electronic Commerce Research and Applications, 8*(1), 48–58.

Executive Order No. 13707, 3 C.F.R. 56365. (2015). https://www.whitehouse.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-american. Accessed 27 Sept 2016.

Federal Trade Commission. (1998). *Privacy online: A report to Congress*. pp. 10–11. Washington, DC. https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-report-congress/priv-23a.pdf. Accessed 27 Sept 2016.

Groom, V., & Calo, M. R. (2011). Reversing the privacy paradox: an experimental study. *TPRC Conference proceedings.* http://ssrn.com/abstract=1993125. Accessed 27 Sept 2016.

Heckman, C. E. & Wobbrock, J. O. (2000). Put your best face forward: Anthropomorphic agents, e-commerce consumers, and the law. In *Proceedings of the fourth international conference on Autonomous agents* (pp. 435–442). New York: ACM.

Hoofnagle, C. J. & King, J. (2008). What Californians understand about privacy online. *SSRN 1262130*. http://dx.doi.org/10.2139/ssrn.1262130. Accessed 27 Sept 2016.

Hoofnagle, C. J., King, J., Li, S., & Turow, J. (2010). How different are young adults from older adults when it comes to information privacy attitudes and policies? http://dx.doi.org/10.2139/ssrn.1589864. Accessed 27 Sept 2016.

Jacoby, J., Speller, D. E., Kohn, C. A. (1974). Brand choice behavior as a function of information load. *Journal of Marketing Research*, *11*(1), 63–69.

John, L. K., Acquisti, A. & Loewenstein, G. (2009). The best of strangers: Context dependent willingness to divulge personal information. *SSRN 1430482*. http://dx.doi.org/10.2139/ssrn.1430482. Accessed 27 Sept 2016.

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity. *European Journal of Social Psychology, 31*(2), 177–192.

Joinson, A. N. & Paine, C. B. (2007). Self-disclosure, privacy and the Internet. *Oxford Handbook of Internet Psychology,* 237-252.

Kahneman, D. (2011). *Thinking, fast and slow.* Penguin.

Lunn, P. (2014). *Regulatory policy and behavioural economics*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264207851-en. Accessed 27 Sept 2016.

Lusoli, W., Bacigalupo, M., Lupiáñez-Villanueva, F., de Andrade, N. N. G., Monteleone, S., Maghiros, I. (2012). *Pan-European survey of practices, attitudes and policy preferences as regards personal identity data management. JRC Scientific and Policy Reports, EUR 25295*. Luxemburg: Luxemburg Publications Office.

Martin, K. (2015). Privacy notices as tabula rasa: an empirical investigation into how complying with a privacy notice is related to meeting privacy expectations online. *Journal of Public Policy and Marketing, 34*(2), 210–227.

Moon, Y. (2000). Intimate exchanges: using computers to elicit self-disclosure from consumers. *Journal of Consumer Research, 26*(4), 323–339.

Nass, C., Kin, E. Y., & Lee, E. J. (1998). *When my face is the interface: An experimental comparison of interacting with one's own face or someone else's face.* In Proceedings of the SIGCHI Conference on Human factors in Computing Systems, Los Angeles.

Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing, 19*(1), 27–41.

Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. *Journal of Management Information Systems, 25*(4), 145–182.

Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing and Management, 44*(1), 386–399.

Solove, D. J. (2012). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, *126*, 1880–1903.

Sousa Lourenço, J., Ciriolo, E., Rafael Almeida, S., & Troussard, X. (2016). Behavioural insights applied to policy: European report 2016. *EUR 27726 EN,*. doi:10.2760/707591.

Steinfeld, N. (2016). "I agree to the terms and conditions":(How) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior, 55*, 992–1000.

Sunstein, C. R. (2014). Nudging: A very short guide. *Journal of Consumer Policy*, *37*(4), 583–588.

Tang, Z., Hu, Y., & Smith, M. D. (2008). Gaining trust through online privacy protection: self-regulation, mandatory standards, or caveat emptor. *Journal of Management Information Systems, 24*(4), 153–173.

Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.

van Bavel, R., Herrmann, B., Esposito, G., & Proestakis, A. (2013). *Applying Behavioural sciences to EU policy-making, JRC Scientific and Policy Reports EUR 26033 EN.* http://ec.europa.eu/dgs/health_consumer/information_sources/docs/30092013_jrc_scientific_policy_report_en.pdf. Accessed 27 Sept 2016.

Wang, Y., Leon, P. G., Scott, K., Chen, X., Acquisti, A. & Cranor, L. F. (2013). Privacy nudges for social media: an exploratory Facebook study. In *Proceedings of the 22nd international conference on World Wide Web companion* (pp. 763–770). International World Wide Web Conferences Steering Committee. New York: ACM.

World Bank (2015). *World Development Report 2015: Mind, Society, and Behaviour* (http://www.worldbank.org/en/publication/wdr2015). Accessed 27 Sept 2016.

Wu, K. W., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior*, *28*(3), 889-897.